Lectures on the Entropy Theory of Measure-Preserving Transformations Part 2

Jonathan Homann
Based on V.A. Rokhlin's Lectures

North West University

1 December 2022

Outline I

Table of Contents

Purpose

Part 1, presented by M. Masmoudi, dealt with the introduction to a measure preserving transformation (including the definition of a Lebesgue space), and concluded with mixing of a measure preserving transformation. The purpose of Part 2 is to introduce the concept of the entropy of a measure preserving transformation.

Table of Contents

Measurable Partition

Let $(X, \mathcal{B}, \mu)$ be a measure space, then $\mathcal{P}$ is a countable measurable partition of $\mathcal{B}$, or, simply, a partition of $\mathcal{B}$, whenever $\mathcal{P}$ is a countable collection of non-empty, pairwise disjoint members of $\mathcal{B}$ which with union X. Furthermore, if $\mathcal{P}$ is a subpartition of some partition of $\mathcal{B}$, $\mathcal{P}_1$, then write $\mathcal{P}_1 \leq \mathcal{P}$, where $\mathcal{P} \geq \mathcal{P}_1$ if and only if $\mathcal{P}_1 \leq \mathcal{P}$.

Product and Intersection Measurable Partitions

Let $(X, \mathcal{B}, \mu)$ be a measure space and let $\{\mathcal{P}_a \mid a \in A\}$ be a family of partitions of $\mathcal{B}$, then the product, denoted $\bigvee_{a \in A} \mathcal{P}_a$, is the partition $\mathcal{P}$ of $\mathcal{B}$ such that $\mathcal{P}_a \leq \mathcal{P}$ for all $a \in A$ and if $\mathcal{P}_a \leq \mathcal{P}'$ for all $a \in A$, for some partition $\mathcal{P}'$, then $\mathcal{P} \leq \mathcal{P}'$.

Product and Intersection Measurable Partitions

Let $(X, \mathcal{B}, \mu)$ be a measure space and let $\{\mathcal{P}_a \mid a \in A\}$ be a family of partitions of $\mathcal{B}$, then the product, denoted $\bigvee_{a\in A} \mathcal{P}_a$, is the partition $\mathcal{P}$ of $\mathcal{B}$ such that $\mathcal{P}_a \leq \mathcal{P}$ for all $a \in A$ and if $\mathcal{P}_a \leq \mathcal{P}'$ for all $a \in A$, for some partition $\mathcal{P}'$, then $\mathcal{P} \leq \mathcal{P}'$. If A is finite, say, $A = \{1, 2, ..., n\}$, then the product may be denoted

$$\bigvee_{i=1}^{n} \mathcal{P}_i = \mathcal{P}_1 \mathcal{P}_2 ... \mathcal{P}_n.$$

Product and Intersection Measurable Partitions

Let $(X, \mathcal{B}, \mu)$ be a measure space and let $\{\mathcal{P}_a \mid a \in A\}$ be a family of partitions of $\mathcal{B}$, then the product, denoted $\bigvee_{a \in A} \mathcal{P}_a$, is the partition $\mathcal{P}$ of $\mathcal{B}$ such that $\mathcal{P}_a \leq \mathcal{P}$ for all $a \in A$ and if $\mathcal{P}_a \leq \mathcal{P}'$ for all $a \in A$, for some partition $\mathcal{P}'$, then $\mathcal{P} \leq \mathcal{P}'$. If A is finite, say, $A = \{1, 2, ..., n\}$, then the product may be denoted

$$\bigvee_{i=1}^{n} \mathcal{P}_i = \mathcal{P}_1 \mathcal{P}_2 ... \mathcal{P}_n.$$

The intersection is denoted $\bigwedge_{a \in A} \mathcal{P}_a$, where it is the partition $\mathcal{P}$ such that $\mathcal{P} \leq \mathcal{P}_a$ for all $a \in A$ and if $\mathcal{P}'$ is a partition of $\mathcal{B}$ such that $\mathcal{P}' \leq \mathcal{P}_a$ for all $a \in A$, then $\mathcal{P}' \leq \mathcal{P}$.

Product and Intersection Measurable Partitions

Let $(X, \mathcal{B}, \mu)$ be a measure space and let $\{\mathcal{P}_a \mid a \in A\}$ be a family of partitions of $\mathcal{B}$, then the product, denoted $\bigvee_{a\in A} \mathcal{P}_a$, is the partition $\mathcal{P}$ of $\mathcal{B}$ such that $\mathcal{P}_a \leq \mathcal{P}$ for all $a \in A$ and if $\mathcal{P}_a \leq \mathcal{P}'$ for all $a \in A$, for some partition $\mathcal{P}'$, then $\mathcal{P} \leq \mathcal{P}'$. If A is finite, say, $A = \{1, 2, ..., n\}$, then the product may be denoted

$$\bigvee_{i=1}^{n} \mathcal{P}_i = \mathcal{P}_1\mathcal{P}_2...\mathcal{P}_n.$$

The intersection is denoted $\bigwedge_{a\in A} \mathcal{P}_a$, where it is the partition $\mathcal{P}$ such that $\mathcal{P} \leq \mathcal{P}_a$ for all $a \in A$ and if $\mathcal{P}'$ is a partition of $\mathcal{B}$ such that $\mathcal{P}' \leq \mathcal{P}_a$ for all $a \in A$, then $\mathcal{P}' \leq \mathcal{P}$.

The symbol $\mathcal{P}_n \nearrow \mathcal{P}$ indicates that $\mathcal{P}_1 \leq \mathcal{P}_2 \leq ...$ and $\mathcal{P} = \bigvee_{n=1}^{\infty} \mathcal{P}_n$.

Product and Intersection Measurable Partitions

Let $(X, \mathcal{B}, \mu)$ be a measure space and let $\{\mathcal{P}_a \mid a \in A\}$ be a family of partitions of $\mathcal{B}$, then the product, denoted $\bigvee_{a \in A} \mathcal{P}_a$, is the partition $\mathcal{P}$ of $\mathcal{B}$ such that $\mathcal{P}_a \leq \mathcal{P}$ for all $a \in A$ and if $\mathcal{P}_a \leq \mathcal{P}'$ for all $a \in A$, for some partition $\mathcal{P}'$, then $\mathcal{P} \leq \mathcal{P}'$. If A is finite, say, $A = \{1, 2, ..., n\}$, then the product may be denoted

$$\bigvee_{i=1}^{n} \mathcal{P}_i = \mathcal{P}_1 \mathcal{P}_2 ... \mathcal{P}_n.$$

The intersection is denoted $\bigwedge_{a \in A} \mathcal{P}_a$, where it is the partition $\mathcal{P}$ such that $\mathcal{P} \leq \mathcal{P}_a$ for all $a \in A$ and if $\mathcal{P}'$ is a partition of $\mathcal{B}$ such that $\mathcal{P}' \leq \mathcal{P}_a$ for all $a \in A$, then $\mathcal{P}' \leq \mathcal{P}$.

The symbol $\mathcal{P}_n \nearrow \mathcal{P}$ indicates that $\mathcal{P}_1 \leq \mathcal{P}_2 \leq ...$ and $\mathcal{P} = \bigvee_{n=1}^{\infty} \mathcal{P}_n$. The symbol $\mathcal{P}_n \searrow \mathcal{P}$ indicates that $\mathcal{P}_1 \geq \mathcal{P}_2 \geq ...$ and $\bigwedge_{n=1}^{\infty} \mathcal{P}_n = \mathcal{P}$.

Table of Contents

Measure Preserving Transformation

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T\colon X \to X$ such that $T^{-1}(\mathcal{B}) \subseteq \mathcal{B}$ and $\mu\left(T^{-1}(A)\right) = \mu(A)$ for each $A \in \mathcal{B}$, then T is called a measure preserving transformation and $(X, \mathcal{B}, \mu, T)$ is called a measure preserving system.

Measure Preserving Transformation

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T\colon X \to X$ such that $T^{-1}(\mathcal{B}) \subseteq \mathcal{B}$ and $\mu\left(T^{-1}(A)\right) = \mu(A)$ for each $A \in \mathcal{B}$, then T is called a measure preserving transformation and $(X, \mathcal{B}, \mu, T)$ is called a measure preserving system. Henceforth, it is assumed that all probability spaces are Lebesgue spaces and, consequently, all measure preserving systems are assumed to be a measure preserving transformation defined on a Lebesgue space with measure 1.

Table of Contents

Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system, let $\mathcal{P}$ be a partition of $\mathcal{B}$ and let $C_1, C_2, \ldots$ be members of $\mathcal{P}$ with strictly positive measure, then the entropy of $\mathcal{P}$ is denoted $\mathrm{H}(\mathcal{P})$, where

$$\mathrm{H}(\mathcal{P}) := \begin{cases} -\sum_n \mu(C_n) \log_2(\mu(C_n)) & \text{if } \mu(X \setminus \bigcup_n C_n) = 0, \\ \infty & \text{if } \mu(X \setminus \bigcup_n C_n) > 0, \end{cases}$$

where $0 \log_2(0) := 0$.

Entropy II

If $m\left(x;\mathcal{P}\right)$ denotes the measure of the element of $\mathcal{P}$ which contains the point $x \in X$ and, with the convention $\log_2\left(0\right) := -\infty$, then

$$\mathrm{H}\left(\mathcal{P}\right) = -\int \log_2\left(m\left(x;\mathcal{P}\right)\right)\,\mathrm{d}\mu$$

Table of Contents

Properties of Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let $\mathcal{P}$ be a partition of $\mathcal{B}$, then

Properties of Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let $\mathcal{P}$ be a partition of $\mathcal{B}$, then

1. $\mathrm{H}(\mathcal{P}) \geq 0$, where $\mathrm{H}(\mathcal{P}) = 0$ if and only if $\mathcal{P} = \{X\}$.

Properties of Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let $\mathcal{P}$ be a partition of $\mathcal{B}$, then

1. $\mathrm{H}(\mathcal{P}) \geq 0$, where $\mathrm{H}(\mathcal{P}) = 0$ if and only if $\mathcal{P} = \{X\}$.
2. If Q is a partition of $\mathcal{B}$ and if $\mathcal{P} \leq Q$, then $\mathrm{H}(\mathcal{P}) \leq \mathrm{H}(Q)$. Furthermore, if $\mathrm{H}(\mathcal{P}) = \mathrm{H}(Q) < \infty$, then $\mathcal{P} = Q$.

Properties of Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let $\mathcal{P}$ be a partition of $\mathcal{B}$, then

1. $\mathrm{H}(\mathcal{P}) \geq 0$, where $\mathrm{H}(\mathcal{P}) = 0$ if and only if $\mathcal{P} = \{X\}$.
2. If Q is a partition of $\mathcal{B}$ and if $\mathcal{P} \leq Q$, then $\mathrm{H}(\mathcal{P}) \leq \mathrm{H}(Q)$. Furthermore, if $\mathrm{H}(\mathcal{P}) = \mathrm{H}(Q) < \infty$, then $\mathcal{P} = Q$.
3. If $\mathcal{P}_1, \mathcal{P}_2, \ldots$ are partitions of $\mathcal{B}$ and $\mathcal{P}_n \nearrow \mathcal{P}$ (resp., $\mathcal{P}_n \searrow \mathcal{P}$), then $\mathrm{H}(\mathcal{P}_n) \nearrow \mathrm{H}(\mathcal{P})$ (resp., $\mathrm{H}(\mathcal{P}_n) \searrow \mathrm{H}(\mathcal{P})$).

Properties of Entropy II

1. If $\mathcal{P}$ has n sets, then $\mathrm{H}(\mathcal{P}) \leq \log_2(n)$. Furthermore, $\mathrm{H}(\mathcal{P}) = \log_2(n)$ if and only if $\mu(P) = \frac{1}{n}$ for each $P \in \mathcal{P}$.

Properties of Entropy II

1. If $\mathcal{P}$ has n sets, then $\mathrm{H}(\mathcal{P}) \leq \log_2(n)$. Furthermore, $\mathrm{H}(\mathcal{P}) = \log_2(n)$ if and only if $\mu(P) = \dfrac{1}{n}$ for each $P \in \mathcal{P}$.
2. If Q is a partition of $\mathcal{B}$, then $\mathrm{H}(\mathcal{P}Q) \leq \mathrm{H}(\mathcal{P}) + \mathrm{H}(Q)$. Furthermore, if $\mathrm{H}(\mathcal{P}), \mathrm{H}(Q) < \infty$, then $\mathrm{H}(\mathcal{P}) = \mathrm{H}(Q)$ if and only if $\mathcal{P}$ and Q are independent (that is, $\mu(A \cap B) = \mu(A)\,\mu(B)$ for each $A \in \mathcal{P}$ and for each $B \in Q$).

Properties of Entropy II

1. If $\mathcal{P}$ has n sets, then $\mathrm{H}(\mathcal{P}) \leq \log_2(n)$. Furthermore, $\mathrm{H}(\mathcal{P}) = \log_2(n)$ if and only if $\mu(P) = \frac{1}{n}$ for each $P \in \mathcal{P}$.
2. If Q is a partition of $\mathcal{B}$, then $\mathrm{H}(\mathcal{P}Q) \leq \mathrm{H}(\mathcal{P}) + \mathrm{H}(Q)$. Furthermore, if $\mathrm{H}(\mathcal{P}), \mathrm{H}(Q) < \infty$, then $\mathrm{H}(\mathcal{P}) = \mathrm{H}(Q)$ if and only if $\mathcal{P}$ and Q are independent (that is, $\mu(A \cap B) = \mu(A)\mu(B)$ for each $A \in \mathcal{P}$ and for each $B \in Q$).
3. If $P_1, \ldots$ is a sequence (finite or infinite), of partitions of $\mathcal{B}$, then $\mathrm{H}\left(\bigvee_n P_n\right) \leq \sum_n \mathrm{H}(P_n)$.

Table of Contents

Opening Observation

If $(X, \mathcal{B}, \mu, T)$ is a measure preserving system and if P and Q are partitions of $\mathcal{B}$, then almost every partition P_B, for $B \in X/Q$, has a well-defined entropy, $\mathrm{H}(P_B)$.

[1] As per the terminology of Riesz spaces, positive includes zero.

Opening Observation

If $(X, \mathcal{B}, \mu, T)$ is a measure preserving system and if P and Q are partitions of $\mathcal{B}$, then almost every partition P_B, for $B \in X/Q$, has a well-defined entropy, $\mathrm{H}(P_B)$.

Notice that $\mathrm{H}(P_B)$ is a positive, measurable function of the factor space X/Q and it is called the conditional entropy of P with respect to Q.

[1] As per the terminology of Riesz spaces, positive includes zero.

Table of Contents

Mean Conditional Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let P and Q be partitions of $\mathcal{B}$, then the mean conditional entropy of P with respect to Q is denoted $\mathrm{H}\left(P/Q\right)$, where

$$\mathrm{H}\left(P/Q\right) := \int_{X/Q} \mathrm{H}\left(P_B\right)\, \mathrm{d}\mu_Q,$$

where μ_Q is the measure defined by $\mu_Q = \mu \circ \rho$, where $\rho \colon X \to Q$, taking x in X to the member of Q in which it is contained.

Mean Conditional Entropy II

An equivalent definition can be sated by letting $B(x)$ being the member of Q which contains $x \in X$ and by letting $m(x; P/Q)$ denote the measure (in $B(x)$), of the member of the partition $P_{B(x)}$ containing x, then

$$\mathrm{H}(P \mid Q) = -\int \log_2 (m(x; P/Q))\, \mathrm{d}\mu.$$

Notice that by using this definition, the domain of integration is X, which is an advantage over the previous definition which has its domain of integration being X/Q.

Table of Contents

Properties of the Mean Conditional Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let P be a partition of $\mathcal{B}$.

1. $\mathrm{H}\left(P / \{X\}\right) = \mathrm{H}\left(P\right)$.

Properties of the Mean Conditional Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let P be a partition of $\mathcal{B}$.

1. $\mathrm{H}(P/\{X\}) = \mathrm{H}(P)$.
2. If Q and R are partitions of $\mathcal{B}$ and if $Q \leq R$, then $\mathrm{H}(PQ/R) = \mathrm{H}(P/R)$.

Properties of the Mean Conditional Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let P be a partition of $\mathcal{B}$.

1. $\mathrm{H}\left(P/\left\{X\right\}\right) = \mathrm{H}\left(P\right)$.
2. If Q and R are partitions of $\mathcal{B}$ and if $Q \leq R$, then $\mathrm{H}\left(PQ/R\right) = \mathrm{H}\left(P/R\right)$.
3. If Q is a partition of $\mathcal{B}$, then $\mathrm{H}\left(P/Q\right) \geq 0$, where $\mathrm{H}\left(P/Q\right) = 0$ if and only if $P \leq Q$.

Properties of the Mean Conditional Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let P be a partition of $\mathcal{B}$.

1. $\mathrm{H}(P/\{X\}) = \mathrm{H}(P)$.
2. If Q and R are partitions of $\mathcal{B}$ and if $Q \leq R$, then $\mathrm{H}(PQ/R) = \mathrm{H}(P/R)$.
3. If Q is a partition of $\mathcal{B}$, then $\mathrm{H}(P/Q) \geq 0$, where $\mathrm{H}(P/Q) = 0$ if and only if $P \leq Q$.
4. If Q and R are partitions of $\mathcal{B}$, then $\mathrm{H}(PQ/R) \leq \mathrm{H}(P/R) + \mathrm{H}(Q/R)$. Furthermore, if $\mathrm{H}(P/R), \mathrm{H}(Q/R) < \infty$, then $\mathrm{H}(PQ/R) = \mathrm{H}(P/R) + \mathrm{H}(Q/R)$ if and only if P and Q are independent with respect to R.

Properties of the Mean Conditional Entropy I

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let P be a partition of $\mathcal{B}$.

1. $\mathrm{H}(P/\{X\}) = \mathrm{H}(P)$.
2. If Q and R are partitions of $\mathcal{B}$ and if $Q \leq R$, then $\mathrm{H}(PQ/R) = \mathrm{H}(P/R)$.
3. If Q is a partition of $\mathcal{B}$, then $\mathrm{H}(P/Q) \geq 0$, where $\mathrm{H}(P/Q) = 0$ if and only if $P \leq Q$.
4. If Q and R are partitions of $\mathcal{B}$, then $\mathrm{H}(PQ/R) \leq \mathrm{H}(P/R) + \mathrm{H}(Q/R)$. Furthermore, if $\mathrm{H}(P/R), \mathrm{H}(Q/R) < \infty$, then $\mathrm{H}(PQ/R) = \mathrm{H}(P/R) + \mathrm{H}(Q/R)$ if and only if P and Q are independent with respect to R.
5. If (P_n) is a sequence of partitions of $\mathcal{B}$ with $P_n \nearrow P$ (resp., $P_n \searrow P$), and if Q is a partition of $\mathcal{B}$, then $\mathrm{H}(P_n/Q) \nearrow \mathrm{H}(P/Q)$ (resp., $\mathrm{H}(P_n/Q) \searrow \mathrm{H}(P/Q)$).

Properties of the Mean Conditional Entropy II

If Q and R are partitions of $\mathcal{B}$, then

$$\mathrm{H}\left(PQ/R\right) = \mathrm{H}\left(P/R\right) + \mathrm{H}\left(Q/PR\right).$$

Opening Statement

In the preceding section, a measure preserving system was used, but only a Lebesgue space (with probability measure), was required. This was no mistake: this was used to enforce the idea that we want to build something which can be applied to a measure preserving transformation on a Lebesgue probability space.

Entropy of a Measure Preserving Transformation

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let P be a partition of $\mathcal{B}$, then the entropy of T with respect to P is denoted $h(T, P)$, where

$$h(T, P) = \mathrm{H}\left(P / T^{-1} P^{-}\right),$$

where $P^{-} := \bigvee_{n=0}^{\infty} T^{-n} P$.

Entropy of a Measure Preserving Transformation

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let P be a partition of $\mathcal{B}$, then the entropy of T with respect to P is denoted $h(T, P)$, where

$$h(T, P) = \mathrm{H}\left(P/T^{-1}P^{-}\right),$$

where $P^{-} := \bigvee_{n=0}^{\infty} T^{-n}P$.

Theorem

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let P and Q be partitions of $\mathcal{B}$. If $P \leq Q$ and if $\mathrm{H}\left(Q/T^{-1}P^{-}\right) < \infty$, then

$$\frac{1}{n}\mathrm{H}\left(\bigvee_{k=0}^{n-1} T^{-k}P/T^{-n}Q^{-}\right) \to h(T, P) \quad \text{as } n \to \infty.$$

References

ROKHLIN, V.A., *Lectures on the Entropy Theory of Measure-Preserving Transformations*, Russ. Math. Surv. **22** 1, 1967.